

6/1
504

REC'D 10 DEC 1999
WIPO PCT

PCT/JP 99/03492

日本国特許庁

PATENT OFFICE
JAPANESE GOVERNMENT

29.06.99

JP99/3492

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

E. J. U.

出願年月日

Date of Application:

1998年 6月30日

出願番号

Application Number:

平成10年特許願第185179号

出願人

Applicant(s):

日本電気株式会社

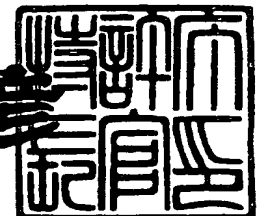
**PRIORITY
DOCUMENT**

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

1999年11月26日

特許庁長官
Commissioner,
Patent Office

近藤 隆彦



出証番号 出証特平11-3081573

【書類名】 特許願

【整理番号】 33509206

【提出日】 平成10年 6月30日

【あて先】 特許庁長官 殿

【国際特許分類】 G10L 9/14
G10L 9/18
H03M 7/30

【発明の名称】 音声符号化装置

【請求項の数】 11

【発明者】
【住所又は居所】 東京都港区芝五丁目7番1号 日本電気株式会社内
【氏名】 小澤 一範

【特許出願人】
【識別番号】 000004237
【氏名又は名称】 日本電気株式会社
【代表者】 金子 尚志

【代理人】
【識別番号】 100080816
【弁理士】
【氏名又は名称】 加藤 朝道
【電話番号】 045-476-1131

【手数料の表示】
【予納台帳番号】 030362
【納付金額】 21,000円

【提出物件の目録】
【物件名】 明細書 1
【物件名】 図面 1
【物件名】 要約書 1
【包括委任状番号】 9304371

特平 10-185179

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声符号化装置

【特許請求の範囲】

【請求項 1】

音声信号を入力しスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、

過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック部と、

前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部と、

を少なくとも含む音声符号化装置において、

前記音声信号から特徴を抽出してモードを判別する判別部と、

前記判別部の出力が予め定められたモードである場合に音源信号を複数の非零のパルスの組合せで表わし、前記パルスの振幅もしくは極性をまとめて量子化するコードブックを有し、

前記コードブックに格納されるコードベクトルと前記パルスの位置をシフトする複数のシフト量とを組み合わせで探索し、入力音声との歪みを最小にするコードベクトルとシフト量の組合せを出力する音源量子化部と、

前記スペクトルパラメータ計算部の出力と前記判別部の出力と前記適応コードブック部の出力と前記音源量子化部の出力とを組み合わせで出力するマルチプレクサ部と、を備えたことを特徴とする音声符号化装置。

【請求項 2】

音声信号を入力しスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、

過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック部と、

前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部と、

を少なくとも含む音声符号化装置において、

前記音声信号から特徴を抽出してモードを判別する判別部と、

前記判別部の出力が予め定められたモードである場合に、音源信号を複数の非零のパルスの組合せで表わし、前記パルスの振幅もしくは極性をまとめて量子化するコードブックを有し、

予め定められた規則により前記パルスの位置を発生し入力音声との歪みを最小にするコードベクトルを出力する音源量子化部と、

前記スペクトルパラメータ計算部の出力と前記判別部の出力と前記適応コードブック部の出力と前記音源量子化部の出力とを組み合わせ出力するマルチプレクサ部と、

を有することを特徴とする音声符号化装置。

【請求項 3】

音声信号を入力しスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、

過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック部と、

前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部と、

を少なくとも有する音声符号化装置において、

前記音声信号から特徴を抽出してモードを判別する判別部と、

前記判別部の出力が予め定められたモードである場合に音源信号を複数の非零のパルスの組合せで表わし、前記パルスの振幅もしくは極性をまとめて量子化するコードブックとゲインを量子化するゲインコードブックを有し、

前記コードブックに格納されるコードベクトルと、前記パルスの位置をシフトする複数のシフト量と、前記ゲインコードブックに格納されるゲインコードベクトルとを組み合わせ探索し、入力音声との歪みを最小にするコードベクトルとシフト量とゲインコードベクトルの組合せを出力する音源量子化部と、

前記スペクトルパラメータ計算部の出力と、前記判別部の出力と、前記適応コードブック部の出力と、前記音源量子化部の出力と、を組み合わせ出力するマルチプレクサ部と、

を有することを特徴とする音声符号化装置。

【請求項 4】

音声信号を入力しスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、

過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック部と、

前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部と、

を少なくとも含む音声符号化装置において、

前記音声信号から特徴を抽出してモードを判別する判別部と、

前記判別部の出力が予め定められたモードである場合に音源信号を複数の非零のパルスの組合せで表わし、前記パルスの振幅もしくは極性をまとめて量子化するコードブックとゲインを量子化するゲインコードブックを有し、

予め定められた規則により前記パルスの位置を発生し入力音声との歪みを最小にするコードベクトルとゲインコードベクトルの組合せを出力する音源量子化部と、

前記スペクトルパラメータ計算部の出力と前記前記判別部の出力と前記適応コードブック部の出力と前記音源量子化部の出力とを組み合わせて出力するマルチプレクサ部と、

を有することを特徴とする音声符号化装置。

【請求項 5】

音声信号を入力しスペクトルパラメータを求めて量子化するスペクトルパラメータ計算手段と、

過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック手段と、

前記音声信号からその特徴量を抽出し有声・無声等に関するモード判別を行なうモード判別手段と、

を含み、さらに、

前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力

する音源量子化手段であって、予め定められたモードの場合に、複数のパルスの振幅もしくは極性をまとめて量子化するコードブックに格納されるコードベクトルと、予め定められたパルスの位置を時間的にシフトする複数のシフト量の各々とを組み合わせ探索し、入力音声との歪みを最小にするコードベクトルのインデクスとシフト量の組合せを選択する音源量子化手段と、

ゲインコードブックを用いてゲインを量子化するゲイン量子化手段と、

前記スペクトルパラメータ計算手段と前記適応コードブック手段と前記音源量子化手段と前記ゲイン量子化手段の各出力とを組み合わせ出力するマルチプレクス手段と、を含むことを特徴とする音声符号化装置。

【請求項 6】

前記音源量子化手段において、前記モード判別手段での判別が予め定められたモードの場合に、パルスの位置として、予め定められた規則に従い発生した位置を用いることを特徴とする請求項 5 記載の音声符号化装置。

【請求項 7】

前記モード判別手段での判別が予め定められたモードの場合に、予め定められた個数のパルスの位置を乱数発生手段で発生し、前記発生された個数の位置を前記音源量子化手段に出力する、ことを特徴とする請求項 5 記載の音声符号化装置。

。

【請求項 8】

前記音源量子化手段において、前記モード判別手段での判別が予め定められたモードの場合に、前記コードブックの全てのコードベクトルとパルスの位置のシフト量との全ての組合せについて所定の歪み量を最小にする順に複数の組合せを選択してゲイン量子化手段へ出力し、

前記ゲイン量子化手段は、前記音源量子化手段からの複数セットの出力の各々に対し、ゲインコードブックを用いてゲインを量子化し、所定の歪み量を最小化するシフト量、音源コードベクトル、ゲインコードベクトルの組合せを選択する、ことを特徴とする請求項 5 記載の音声符号化装置。

【請求項 9】

前記モード判別手段が、音声信号の前記特徴量としてピッチ予測ゲインを用い

、サブフレーム毎に求めたピッチ予測ゲインの値と、予め定められたしきい値とを比較し、前記ピッチ予測ゲインが前記しきい値より大であれば有声、しきい値未満であれば無声と判断する、ことを特徴とする請求項5記載の音声符号化装置。

【請求項10】

前記予め定められたモードが無音であることを特徴とする請求項5乃至9のいずれかに記載の音声符号化装置。

【請求項11】

請求項1乃至10のいずれかに記載の音声符号化装置と、

前記音声符号化装置の符号化出力を入力し、スペクトルパラメータと適応コードブックの遅延と適応コードベクトルと音源のゲインと、音源情報として振幅もしくは極性コードベクトルとパルスの位置を表す符号を分離して出力するデマルチプレクサ手段と、

適応コードブックにおける過去の量子化されたゲインを用いてモードを判別するモード判別手段と、

前記モード判別手段の出力が予め定められたモードの場合に、予め定められた規則によりパルスの位置を発生し、コードベクトルから前記パルスの振幅もしくは極性を発生して音源信号を復元する音源信号復元手段と、

前記音源信号を通して音声信号を再生する合成フィルタ部と、を少なくとも備えたことを特徴とする音声符号化復号化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音声符号化装置に関し、特に音声信号を低いビットレートで高品質に符号化するための音声符号化装置に関する。

【0002】

【従来の技術】

音声信号を高効率に符号化する従来の方式としては、例えば、M.Schroeder及びB.Atal氏による“Code-excited linear prediction; High quality spe

ech at very low bit rates" (Proc. ICASSP, pp.937-940,1985年) と題した論文 (文献1) や、Kleijn氏らによる "Improved speech quality and efficient vector quantization in SELP" (Proc. ICASSP, pp.155-158, 1988年) と題した論文 (文献2) などに記載されている CELP (Code Excited Linear Predictive Coding; 符号励振型線形予測符号化) が知られている。

【0003】

この CELP 符号化方式では、送信側では、フレーム毎 (例えば 20ms) に音声信号から線形予測 (LPC) 分析を用いて、音声信号のスペクトル特性を表すスペクトルパラメータを抽出する。フレームをさらにサブフレーム (例えば 5ms) に分割し、サブフレーム毎に過去の音源信号を基に適応コードブックにおけるパラメータ (ピッチ周期に対応する遅延パラメータとゲインパラメータ) を抽出し、適応コードブックにより前記サブフレームの音声信号をピッチ予測する。

【0004】

ピッチ予測して求めた音源信号に対して、予め定められた種類の雑音信号からなる音源コードブック (ベクトル量子化コードブック) から最適な音源コードベクトルを選択し、最適なゲインを計算することにより、音源信号を量子化する。

【0005】

音源コードベクトルの選択の仕方は、選択した雑音信号により合成した信号と、前記残差信号との誤差電力を最小化するように行う。そして、選択されたコードベクトルの種類を表すインデクスとゲインならびに、前記スペクトルパラメータと適応コードブックのパラメータをマルチプレクサ部により組み合わせて伝送する。なお、受信側の説明は省略する。

【0006】

【発明が解決しようとする課題】

しかしながら、前記従来の符号化方式では、音源コードブックから最適な音源コードベクトルを選択するのに多大な演算量を要するという問題があった。これは、上記文献1や文献2記載の方法では、音源コードベクトルを選択するのに、各コードベクトルに対して、一旦フィルタリングもしくは畳み込み演算を行ない

、この演算をコードブックに格納されているコードベクトルの個数だけ繰り返すことに起因する。例えば、コードブックのビット数がBビットで、次元数がNのときは、フィルタリングあるいは畳み込み演算のときのフィルタあるいはインパルス応答長をKとすると、演算量は1秒当たり、 $N \times K \times 2^B \times 8000 / N$ だけ必要となる。一例として、B=10、N=40、k=10とすると、1秒当たり81,920,000回の演算が必要となり、極めて膨大であるという問題点があった。

【0007】

音源コードブック探索に必要な演算量の低減を図るための方法として、従来より、種々の方法が提案されている。例えば、ACELP (Algebraic Code Excited Linear Prediction) 方式が提案されている。これは、例えば、C. Laflamme氏らによる“16kbps wideband speech coding technique based on algebraic CELP”と題した論文(Proc. ICASSP, pp.13-16, 1991) (文献3)等を参照することができる。

【0008】

上記文献3の方法によれば、音源信号を複数個のパルスで表し、各パルスの位置を予め定められたビット数で表し伝送する。ここで、各パルスの振幅は+1.0もしくは-1.0に限定されているため、パルス探索の演算量を大幅に低減化できる。

【0009】

このように、上記文献3の方法では、演算量を大幅に低減化することが可能となる。

【0010】

また、8kb/s以上のビットレートでは良好な音質が得られるが、それ未満のビットレートでは、特に音声に背景雑音が重畳している場合に、符号化音声の背景雑音部分の音質が極めて劣化する、という問題点があった。

【0011】

この理由としては、音源信号を複数個のパルスの組合せで表すので、音声の母音区間では、パルスがピッチの開始点であるピッチパルスの近辺に集中するため少ない個数のパルスで効率的に表すことができるが、背景雑音のようなランダ

ム信号に対しては、パルスをランダムに立てる必要があるため、少ない個数のパルスでは、背景雑音を良好に表すことは困難であり、ビットレートを低減化し、パルスの個数が削減されると、背景雑音に対する音質が急激に劣化するためである。

【0012】

したがって本発明の目的は、上記問題点を解決し、ビットレートが低い場合にも、比較的少ない演算量で、特に背景雑音に対する音質の劣化の少ない音声符号化装置を提供することにある。

【0013】

【課題を解決するための手段】

前記目的を達成するため、本願第1発明の音声符号化装置は、音声信号を入力しスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部と、を少なくとも含む音声符号化装置において、前記音声信号からその特徴を抽出してモードを判別する判別部と、前記判別部の出力が予め定められたモードである場合に音源信号を複数の非零のパルスの組合せで表わし、前記パルスの振幅もしくは極性をまとめて量子化するコードブックを有し、前記コードブックに格納されるコードベクトルと前記パルスの位置をシフトする複数のシフト量とを組み合わせで探索し、入力音声との歪みを最小にするコードベクトルとシフト量の組合せを出力する音源量子化部と、前記スペクトルパラメータ計算部の出力と、前記判別部の出力と、前記適応コードブック部の出力と、前記音源量子化部の出力と、を組み合わせで出力するマルチプレクサ部と、を備える。

【0014】

本願第2発明の音声符号化装置は、前記音声信号から特徴を抽出してモードを判別する判別部と、前記判別部の出力が予め定められたモードである場合に、音源信号を複数の非零のパルスの組合せで表わし、前記パルスの振幅もしくは極性をまとめて量子化するコードブックを有し、予め定められた規則により前記

パルスの位置を発生し入力音声との歪みを最小にするコードベクトルを出力する音源量子化部と、前記スペクトルパラメータ計算部の出力と前記判別部の出力と前記適応コードブック部の出力と前記音源量子化部の出力とを組み合わせで出力するマルチプレクサ部と、を有する。

【0015】

本願第3発明の音声符号化装置は、前記音声信号から特徴を抽出してモードを判別する判別部と、前記判別部の出力が予め定められたモードである場合に音源信号を複数個の非零のパルスの組合せで表わし、前記パルスの振幅もしくは極性をまとめて量子化するコードブックとゲインを量子化するゲインコードブックを有し、前記コードブックに格納されるコードベクトルと、前記パルスの位置をシフトする複数のシフト量と、前記ゲインコードブックに格納されるゲインコードベクトルとを組み合わせで探索し、入力音声との歪みを最小にするコードベクトルとシフト量とゲインコードベクトルの組合せを出力する音源量子化部と、前記スペクトルパラメータ計算部の出力と、前記判別部の出力と、前記適応コードブック部の出力と、前記音源量子化部の出力と、を組み合わせで出力するマルチプレクサ部と、を有する。

【0016】

本願第4発明の音声符号化装置は、前記音声信号から特徴を抽出してモードを判別する判別部と、前記判別部の出力が予め定められたモードである場合に音源信号を複数個の非零のパルスの組合せで表わし、前記パルスの振幅もしくは極性をまとめて量子化するコードブックとゲインを量子化するゲインコードブックを有し、予め定められた規則により前記パルスの位置を発生し入力音声との歪みを最小にするコードベクトルとゲインコードベクトルの組合せを出力する音源量子化部と、前記スペクトルパラメータ計算部の出力と、前記前記判別部の出力と、前記適応コードブック部の出力と、前記音源量子化部の出力と、を組み合わせで出力するマルチプレクサ部と、を有する。

【0017】

【発明の実施の形態】

本発明の実施の形態について以下に説明する。本発明の音声符号化装置は、そ

の一実施の形態において、モード判別回路（図 1 の 800）において、音声信号から特徴量を抽出し、その特徴量をもとに、モードを判別し、音源量子化回路（図 1 の 350）において、予め定められたモードの場合に、複数のパルスの振幅もしくは極性をまとめて量子化するコードブック（図 1 の 351、352）に格納されるコードベクトルと、予め定められたパルスの位置を時間的にシフトする複数のシフト量の各々とを組み合わせ探索し、入力音声との歪みを最小にするコードベクトルとシフト量の組合せを選択し、ゲイン量子化回路（図 1 の 365）において、ゲインコードブック（図 1 の 380）を用いてゲインを量子化し、スペクトルパラメータ計算部（図 1 の 210）の出力と、モード判別部（図 1 の 800）の出力と、適応コードブック回路（図 1 の 500）の出力と、音源量子化部（図 1 の 350）の出力と、ゲイン量子化回路の出力をマルチプレクサ部（図 1 の 400）で組み合わせて出力する。

【0018】

本発明の音声復号化装置は、その好ましい一実施の形態において、入力端子から入力した符号系列を分離し、スペクトルパラメータと適応コードブックの遅延、適応コードベクトルと音源のゲイン、音源情報として振幅もしくは極性コードベクトルとパルスの位置を表す符号を分離して出力するデマルチプレクサ部（図 5 の 510）と、適応コードブックにおける過去の量子化されたゲインを用いてモードを判別するモード判別部（図 5 の 530）と、モード判別部の出力が予め定められたモードの場合に、量子化された音源情報から非零のパルスを発生して音源信号を復元する音源信号復元部（図 5 の 540）と、を備え、合成フィルタ部（図 5 の 560）に前記音源信号を通して音声信号を再生する。

【0019】

【実施例】

本発明の実施例について図面を参照して以下に説明する。

【0020】

[実施例 1]

図 1 は、本発明の音声符号化装置の一実施例の構成を示すブロック図である。

【0021】

図1を参照すると、入力端子100から音声信号を入力し、フレーム分割回路110では音声信号をフレーム（例えば20ms）毎に分割し、サブフレーム分割回路120では、フレームの音声信号をフレームよりも短いサブフレーム（例えば5ms）に分割する。

【0022】

スペクトルパラメータ計算回路200では、少なくとも一つのサブフレームの音声信号に対して、サブフレーム長よりも長い窓（例えば24ms）をかけて音声を取り出してスペクトルパラメータを予め定められた次数（例えば $P=10$ 次）計算する。ここで、スペクトルパラメータの計算には、周知のLPC（線形予測符号化）分析や、Burg（ブルグ）分析等を用いることができる。ここでは、Burg分析を用いることとする。なお、Burg分析の詳細については、中溝著による「信号解析とシステム同定」と題した刊行物（コロナ社1988年刊）の第82～87頁（文献4）等の記載が参照されるので説明は略する。

【0023】

さらにスペクトルパラメータ計算部210では、Burg法により計算された線形予測係数 α_i ($i=1, \dots, 10$)を量子化や補間に適したLSPパラメータに変換する。ここで、線形予測係数からLSPへの変換は、菅村他による“線スペクトル対（LSP）音声分析合成方式による音声情報圧縮”と題した論文（電子通信学会論文誌、J64-A、pp.599-606、1981年）（文献5）を参照することができる。例えば、第2、4サブフレームでBurg法により求めた線形予測係数を、LSPパラメータに変換し、第1、3サブフレームのLSPを直線補間により求めて、第1、3サブフレームのLSPを逆変換して線形予測係数に戻し、第1-4サブフレームの線形予測係数 α_{il} ($i=1, \dots, 10, l=1, \dots, 5$)を聴感重み付け回路230に出力する。また、第4サブフレームのLSPをスペクトルパラメータ量子化回路210へ出力する。

【0024】

スペクトルパラメータ量子化回路210では、予め定められたサブフレームのLSPパラメータを効率的に量子化し、次式（1）の歪みを最小化する量子化値を出力する。

【0025】

【数1】

$$D_j = \sum_{i=1}^p W(i) [LSP(i) - QLSP(i)_j]^2 \quad \text{----- (1)}$$

【0026】

ここで、LSP (i) , QLSP (i) j、W (i) はそれぞれ、量子化前のi次目のLSP、量子化後のj番目の結果、重み係数である。

【0027】

以下では、量子化法として、ベクトル量子化を用いるものとし、第4サブフレームのLSPパラメータを量子化するものとする。LSPパラメータのベクトル量子化の手法は周知の手法を用いることができる。具体的な方法としては、例えば、特開平4-171500号公報（特願平2-297600号）（文献6）や特開平4-363000号公報（特願平3-261925号）（文献7）や、特開平5-6199号公報（特願平3-155049号）（文献8）や、T.Nomura et al., による“LSP Coding Using VQ-SVQ With Interpolation in 4.075 kbps M-LCELP Speech Coder”と題した論文（Proc. Mobile Multimedia Communications, pp.B.2.5, 1993）（文献9）等を参照できるのでここでは説明は略する。

【0028】

また、スペクトルパラメータ量子化回路210では、第4サブフレームで量子化したLSPパラメータをもとに、第1～第4サブフレームのLSPパラメータを復元する。ここでは、現フレームの第4サブフレームの量子化LSPパラメータと1つ過去のフレームの第4サブフレームの量子化LSPを直線補間して、第1～第3サブフレームのLSPを復元する。ここで、量子化前のLSPと量子化後のLSPとの誤差電力を最小化するコードベクトルを1種類選択した後に、直線補間により第1～第4サブフレームのLSPを復元できる。さらに性能を向上させるためには、前記誤差電力を最小化するコードベクトルを複数候補選択したのちに、各々の候補について、累積歪を評価し、累積歪を最小化する候補と補間LSPの組を選択するようにすることができる。詳細は、例えば、特願平5-8737

号明細書（文献 10）を参照することができる。

【0029】

以上により復元した第1-3サブフレームのLSPと第4サブフレームの量子化LSPをサブフレーム毎に線形予測係数 α_{il} ($i=1, \dots, 10, l=1, \dots, 5$)に変換し、インパルス応答計算回路310へ出力する。また、第4サブフレームの量子化LSPのコードベクトルを表すインデックスをマルチプレクサ400に出力する。

【0030】

聴感重み付け回路230は、スペクトルパラメータ計算回路200から、各サブフレーム毎に量子化前の線形予測係数 α_{il} ($i=1, \dots, 10, l=1, \dots, 5$)を入力し、前記文献1にもとづき、サブフレームの音声信号に対して聴感重み付けを行い、聴感重み付け信号を出力する。

【0031】

応答信号計算回路240は、スペクトルパラメータ計算回路200から、各サブフレーム毎に線形予測係数 α_{il} を入力し、スペクトルパラメータ量子化回路210から、量子化、補間して復元した線形予測係数 α_{il} をサブフレーム毎に入力し、保存されているフィルタメモリの値を用いて、入力信号を零 $d(n)=0$ とした応答信号を1サブフレーム分計算し、減算器235へ出力する。ここで、応答信号 $x_z(n)$ は下式で表される。

【0032】

【数2】

$$x_z(n) = d(n) - \sum_{i=1}^{10} \alpha_i d(n-i) + \sum_{i=1}^{10} \alpha_i r^i y(n-i) + \sum_{i=1}^{10} \alpha_i' r^i x_x(n-i) \quad \text{---(2)}$$

但し、 $n-i \leq 0$ のときは

$$y(n-i) = p(N+(n-i)) \quad \text{---(3)}$$

$$x_z(n-i) = s_x(N+(n-i)) \quad \text{---(4)}$$

【0033】

ここでNはサブフレーム長を示す。 γ は、聴感重み付け量を制御する重み係数であり、下記の式(7)と同一の値である。 $sw(n)$ 、 $p(n)$ は、それぞれ、重み付け信号計算回路の出力信号、後述の式(7)における右辺第1項のフィルタの分母の項の出力信号をそれぞれ示す。

【0034】

減算器235は、次式(5)により、聴感重み付け信号から応答信号を1サブフレーム分減算し、 $x'w(n)$ を適応コードブック回路300へ出力する。

【0035】

【数3】

$$x'_w(n) = x_w(n) - x_x(n) \quad \text{----- (5)}$$

【0036】

インパルス応答計算回路310は、z変換が次式(6)で表される伝達函数 $H_w(z)$ を有する聴感重み付けフィルタのインパルス応答 $hw(n)$ を予め定められた点数しだけ計算し、適応コードブック回路500、音源量子化回路350へ出力する。

【0037】

【数4】

$$H_w(z) = \frac{1 - \sum_{i=1}^{10} \alpha_i z^{-i}}{1 - \sum_{i=1}^{10} \alpha_i \gamma^i z^{-i} - \sum_{i=1}^{10} \alpha'_i \gamma^i z^{-i}} \quad \text{----- (6)}$$

【0038】

モード判別回路800は、サブフレーム分割回路120の出力信号を用いて、特徴量を抽出し、サブフレーム毎に有声か無声かの判別を行う。ここで、特徴と

しては、ピッチ予測ゲインを用いることができる。サブフレーム毎に求めたピッチ予測ゲインの値と予め定められたしきい値を比較し、ピッチ予測ゲインがしきい値より大きければ有声、さもなければ無声と判断する。

【0039】

モード判別回路 800 は、有声・無声判別情報を音源量子化回路 350 とゲイン量子化回路 365 とマルチプレクサ 400 へ出力する。

【0040】

適応コードブック回路 500 では、ゲイン量子化回路 365 から過去の音源信号 $v(n)$ を、減算器 235 から出力信号 $x'_w(n)$ を、インパルス応答計算回路 310 から聴感重み付けインパルス応答 $h_w(n)$ を入力する。ピッチに対応する遅延 T を次式 (7) の歪み D_T を最小化するように求め、遅延を表すインデックスをマルチプレクサ 400 に出力する。

【0041】

【数 5】

$$D_T = \sum_{n=0}^{N-1} x_w'^2(n) - \left[\sum_{n=0}^{N-1} x_w'(n) y_w(n-T) \right]^2 / \left[\sum_{n=0}^{N-1} y_w^2(n-T) \right] \quad \text{----- (7)}$$

$$y_w(n-T) = v(n-T) * h_w(n) \quad \text{----- (8)}$$

【0042】

式 (8) において、記号 * は畳み込み (convolution) 演算を表す。

【0043】

そしてゲイン β を次式 (9) に従い求める。

【0044】

【数 6】

$$\beta = \sum_{n=0}^{N-1} x_w'(n) y_w(n-T) / \sum_{n=0}^{N-1} y_w^2(n-T) \quad \text{----- (9)}$$

【0045】

ここで、女性音や、子供の声に対して、遅延の抽出精度を向上させるために、

遅延を整数サンプルではなく、小数サンプル値で求めてもよい。具体的な方法は、例えば、P.Kroonらによる、“Pitch pre-dictors with high temporal resolution”と題した論文（Proc. ICASSP, pp.661-664, 1990年）（文献11）等を参照することができる。

【0046】

さらに、適応コードブック回路500では次式(10)に従いピッチ予測を行ない、予測残差信号 $e_w(n)$ を音源量子化回路350へ出力する。

【0047】

【数7】

$$e_w(n) = x'_w(n) - \beta v(n-T) * h_w(n) \text{ ----- (10)}$$

【0048】

音源量子化回路350では、モード判別回路800からの有声・無声判別情報を入力し、有声と無声でパルスを切替える。

【0049】

有声では、M個のパルスをたてるとする。

【0050】

有声では、パルスの振幅をMパルス分まとめて量子化するための、Bビットの振幅コードブック、もしくは極性コードブックを有しているものとする。以下では、極性コードブックを用いる場合について説明する。

【0051】

この極性コードブックは、有声の場合は音源コードブック351に、無声の場合は音源コードブック352に格納されている。

【0052】

有声では、音源量子化回路350は、音源コードブック351から極性コードベクトルを読みだし、各コードベクトルに対して位置をあてはめ、次式(11)の D_k を最小化するコードベクトルと位置の組合せを選択する。

【0053】

【数8】

$$D_k = \sum_{n=0}^{N-1} [e_w(n) - \sum_{i=1}^M g'_{ik} h_w(n-m_i)]^2 \quad \text{----- (11)}$$

【0054】

ここで、 $h_w(n)$ は、聴感重み付けインパルス応答である。

【0055】

上式(11)を最小化するには、次式(12)の $D_{(k,i)}$ を最大化する振幅コードベクトル k と位置 m_i の組合せを求めれば良い。

【0056】

【数9】

$$D_{(k,i)} = \left[\sum_{n=0}^{N-1} e_w(n) s_{wk}(m_i) \right]^2 / \sum_{n=0}^{N-1} s_{wk}^2(m_i) \quad \text{----- (12)}$$

【0057】

ここで、 $s_{wk}(m_i)$ は式(5)で計算される。

【0058】

または、次式(13)の $D_{(k,i)}$ を最大化するように選択しても良い。この方が分子に計算に要する演算量が低減化される。

【0059】

【数10】

$$D_{(k,i)} = \left[\sum_{n=0}^{N-1} \phi(n) v_k(n) \right]^2 / \sum_{n=0}^{N-1} s_{wk}^2(m_i) \quad \text{----- (13)}$$

$$\phi(n) = \sum_{i=n}^{N-1} e_w(i) h_w(i-n), n=0, \dots, N-1 \quad \text{----- (14)}$$

【0060】

ここで、有聲の場合の各パルスのとり得る位置は、演算量削減のため、上記文献3に示すように、拘束することができる。一例として、 $N=40$ 、 $M=5$ とすると、各パルスのとり得る位置は、表1のようになる。

【0061】

【表1】

0, 5, 10, 15, 20, 25, 30, 35
1, 6, 11, 16, 21, 26, 31, 36
2, 7, 12, 17, 22, 27, 32, 37
3, 8, 13, 18, 23, 28, 33, 38
4, 9, 14, 19, 24, 29, 34, 39

【0062】

そして、コードベクトルを表すインデックスをマルチプレクサ400に出力する。

【0063】

さらに、パルスの位置を予め定められたビット数で量子化し、位置を表すインデックスをマルチプレクサに出力する。

【0064】

無声では、表2に示すように、パルスの位置を一定の間隔で定め、パルス全体の位置をシフトさせるためのシフト量を定めておく。以下の例の場合は、1サンプルずつシフトさせるとして、4種類のシフト量（シフト0、シフト1、シフト2、シフト3）を用いることができる。また、この場合、シフト量を2ビットで量子化して伝送する。

【0065】

【表2】

パルスの位置

0, 4, 8, 12, 16, 20, 24, 28,

【0066】

さらに、各々のシフト量に対して、極性コードブック352から極性コードベクトルを入力し、全てのシフト量と全てのコードベクトルの組合せ探索を行ない、次式(15)の歪み $D_{k,j}$ を最小化するシフト量 $\delta(j)$ とコードベクトル g_k の組合せを選択する。

【0067】

【数11】

$$D_{k,j} = \sum_{n=0}^{N-1} [e_w(n) - \sum_{i=1}^M g'_{ik} h_w(n-m_i - \delta(j))]^2 \quad \text{----- (15)}$$

【0068】

選択されたコードベクトルを表すインデクスとシフト量を表す符号をマルチプレクサ400に出力する。

【0069】

なお、複数パルスの振幅を量子化するためのコードブックを、音声信号を用いて予め学習して格納しておくようにしてもよい。コードブックの学習法としては、例えば、Linde氏らによる“An algorithm for vector quantization design,”と題した論文(IEEE Trans. Commun., pp.84-95, January, 1980)(文献12)等を参照できる。

【0070】

有声・無声の場合の振幅、位置の情報は、ゲイン量子化回路365に出力される。

【0071】

ゲイン量子化回路365は、音源量子化回路350から振幅、位置情報を入力

し、モード判別回路 800 から有声・無声判別情報を入力する。

【0072】

ゲインコードブック 380 からゲインコードベクトルを読み出し、選択された振幅コードベクトルもしくは極性コードベクトルと位置に対して、次式 (16) の D_K を最小化するようにゲインコードベクトルを選択する。

【0073】

ここでは、適応コードブックのゲインとパルスで表した音源のゲインの両者を同時にベクトル量子化する例について示す。

【0074】

判別情報が有声の場合には、次式 (16) の D_K を最小化するようにゲインコードベクトルを求める。

【0075】

【数 12】

$$D_k = \sum_{n=0}^{N-1} [x_v(n) - \beta'_i v(n-T) * h_w(n) - G'_i \sum_{j=1}^M g'_{ik} h_w(n-m_j)]^2 \quad \text{----- (16)}$$

【0076】

ここで、 β_k 、 G_k は、ゲインコードブック 355 に格納された 2 次元ゲインコードブックにおける k 番目のコードベクトルである。選択されたゲインコードベクトルを表すインデックスをマルチプレクサ 400 に出力する。

【0077】

一方、判別情報が無声の場合には、次式 (17) の D_K を最小化するよう、ゲインコードベクトルを探索する。

【0078】

【数 13】

$$D_k = \sum_{n=0}^{N-1} [x_v(n) - \beta'_i v(n-T) * h_w(n) - G'_i \sum_{j=1}^M g'_{ik} h_w(n-m_j - \delta(j))]^2 \quad \text{---- (17)}$$

【0079】

選択されたゲインコードベクトルを表すインデックスをマルチプレクサ 400 に出力する。

【0080】

重み付け信号計算回路360は、有声・無声判別情報とそれぞれのインデックスを入力し、インデックスからそれに対応するコードベクトルを読み出す。有声の場合は、次式(18)に基づき駆動音源信号 $v(n)$ を求める。

【0081】

【数14】

$$v(n) = \beta'_i v(n-T) + G'_i \sum_{k=1}^N g'_{ik} \delta(n-m_i) \quad \text{----- (18)}$$

【0082】

$v(n)$ は適応コードブック回路500に出力される。

【0083】

無声の場合には、次式(19)に基づき駆動音源信号 $v(n)$ を求める。

【0084】

【数15】

$$v(n) = \beta'_i v(n-T) + G'_i \sum_{k=1}^N g'_{ik} \delta(n-m_i - \delta(j)) \quad \text{----- (19)}$$

【0085】

$v(n)$ は適応コードブック回路500に出力される。

【0086】

次に、スペクトルパラメータ計算回路200の出力パラメータ、スペクトルパラメータ量子化回路210の出力パラメータを用いて次式(20)により、応答信号 $s_w(n)$ をサブフレーム毎に計算し、応答信号計算回路240へ出力する。

【0087】

【数16】

$$s_w(n) = v(n) - \sum_{i=1}^{10} a_i v(n-i) + \sum_{i=1}^{10} a_i \gamma^i p(n-i) + \sum_{i=1}^{10} a_i \gamma^i s_w(n-i) \quad \text{---- (20)}$$

【0088】

[実施例2]

次に本発明の第2の実施例について説明する。図2は、本発明の第2の実施例の構成を示すブロック図である。図2を参照すると、本発明の第2の実施例においては、前記第1の実施例と、音源量子化回路355の動作が相違している。すなわち、本発明の第2の実施例では、有声・無声判別情報が無声の場合に、パルスの位置として、あらかじめ定められた規則に従い発生した位置を使用する。

【0089】

例えば、予め定められた個数（例えばM1）のパルスの位置を乱数発生回路600により発生させる。つまり、乱数発生器600により発生されたM1個の数値をパルスの位置と考える。これにより発生されたM1個の位置を音源量子化回路355に出力する。

【0090】

音源量子化回路355は、判別情報が有声の場合は、図1の音源量子化回路350と同一の動作を行ない、無声の場合は、乱数発生回路600から出力された位置に対して、音源コードブック352を用いてパルスの振幅もしくは極性をまとめて量子化する。

【0091】

[実施例3]

次に、本発明の第3の実施例について説明する。図3は、本発明の第3の実施例の構成を示すブロック図である。図3を参照

音源量子化回路356では、有声・無声判別情報が無声の場合に、音源コードブック352の全てのコードベクトルとパルスの位置のシフト量との全ての組合せについて、次式による歪みを計算し、次式(21)の $D_{k,j}$ を最小にする順に、複数個の組合せを選択し、ゲイン量子化回路366へ出力する。

【0092】

【数17】

$$D_{k,j} = \sum_{n=0}^{N-1} [e_v(n) - \sum_{i=1}^M g'_{ik} h_v(n-m_i - \delta(j))]^2 \text{ ----- (21)}$$

【0093】

ゲイン量子化回路366は、音源量子化回路356に複数セットの出力の各々

に対し、ゲインコードブック 380 を用いてゲインを量子化し、次式 (22) の $D_{k,j}$ を最小化するシフト量、音源コードベクトル、ゲインコードベクトルの組合せを選択する。

【0094】

【数 18】

$$D_{k,j} = \sum_{n=0}^{N-1} [x_w(n) - \beta'_i v(n-T) * h_w(n) - G'_i \sum_{i=1}^M g'_{ik} h_w(n-m_i - \delta(j))]^2 \text{---(22)}$$

【0095】

【実施例 4】

次に本発明の第 4 の実施例について説明する。図 4 は、本発明の第 4 の実施例の構成を示すブロック図である。図 4 を参照すると、音源量子化回路 357 は、有声・無声判別情報が無声の場合に、乱数発生器 600 において発生されたパルスの位置に対して、音源コードブック 352 を用いてパルスの振幅もしくは極性をまとめて量子化し、全てのコードベクトル、もしくはコードベクトルの複数個の候補をゲイン量子化回路 367 に出力する。

【0096】

ゲイン量子化回路 367 では、音源量子化回路 357 から出力された候補の各々に対してゲインコードブック 380 により、ゲインを量子化し、歪みを最小化するコードベクトルとゲインコードベクトルの組合せを出力する。

【0097】

【実施例 5】

次に本発明の第 5 の実施例について説明する。図 5 は、本発明の第 5 の実施例の構成を示すブロック図である。図 5 を参照すると、デマルチプレクサ 510 は、入力端子 500 から入力した符号系列を分離し、スペクトルパラメータと適応コードブックの遅延、適応コードベクトルと音源のゲイン、音源情報として振幅もしくは極性コードベクトルとパルスの位置を表す符号を分離して出力する。

【0098】

ゲイン復号回路 510 は、ゲインコードブック 380 を用いて適応コードブックと音源のゲインを復号して出力する。

【0099】

適応コードブック回路520は、遅延と適応コードベクトルのゲインを復号化し、過去のサブフレームでの合成フィルタ入力信号を用いて適応コードブック再生信号を発生する。

【0100】

モード判別回路530は、過去のサブフレームで復号した適応コードブックゲインを用いて、予め定められたしきい値と比較し、現在のサブフレームが有声か無声かを判別し、有声・無声判別情報を音源信号復元回路540に出力する。

【0101】

音源信号復元回路540は、有声・無声判別情報を入力し、有声のときは、パルスの位置を復号し、音源コードブック351からコードベクトルを読み出して振幅もしくは極性を与えて、サブフレーム当たり定められた個数のパルスを発生させて音源信号を復元する。

【0102】

無声のときは、予め定められたパルスの位置とシフト量と振幅もしくは極性コードベクトルからパルスを発生させて音源信号を復元する。

【0103】

スペクトルパラメータ復号回路570は、スペクトルパラメータを復号し、合成フィルタ回路560へ出力する。

【0104】

加算器550は、適応コードブック出力信号と音源信号復元回路540の出力信号を加算して合成フィルタ回路560へ出力する。

【0105】

合成フィルタ回路560は、加算器550の出力を入力し、音声を再生して端子580から出力する。

【0106】

【発明の効果】

以上説明したように、本発明によれば、適応コードブックの過去の量子化ゲインをもとに、モードを判別し、予め定められたモードの場合に、複数個のパルス

の振幅もしくは極性をまとめて量子化するコードブックに格納されるコードベクトルと予め定められたパルスの位置を時間的にシフトする複数のシフト量の各々とを組み合わせて探索し、入力音声との歪みを最小にするコードベクトルとシフト量の組合せを選択する構成としたことにより、ビットレートが低い場合にも、比較的少ない演算量で、背景雑音部分が良好に符号化することができる、という顕著な効果を奏する。

【0107】

また、本発明によれば、コードベクトルと複数のシフト量の各々とゲインを量子化するためのゲインコードブックに格納されるゲインコードベクトルとを組み合わせて探索し、入力音声との歪みを最小にするコードベクトルとシフト量とゲインコードベクトルの組合せを選択しているのので、背景雑音が重畳した音声を低ビットレートで符号化しても、背景雑音部分が良好に符号化することができる、という顕著な効果を奏する。

【図面の簡単な説明】

【図1】

本発明の第1の実施例の構成を示すブロック図である。

【図2】

本発明の第2の実施例の構成を示すブロック図である。

【図3】

本発明の第3の実施例の構成を示すブロック図である。

【図4】

本発明の第4の実施例の構成を示すブロック図である。

【図5】

本発明の第5の実施例の構成を示すブロック図である。

【符号の説明】

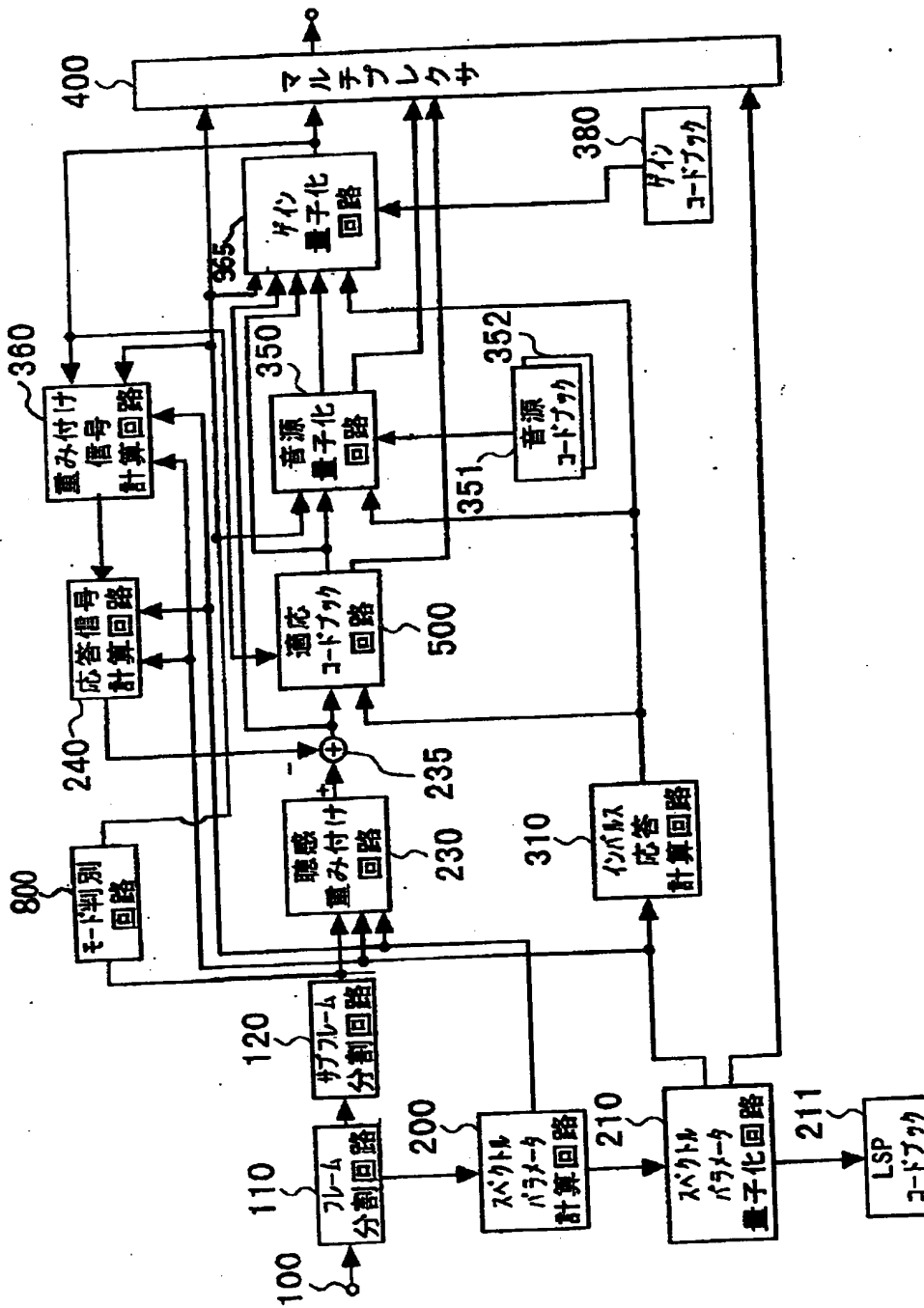
- 110 フレーム分割回路
- 120 サブフレーム分割回路
- 200 スペクトルパラメータ計算回路
- 210 スペクトルパラメータ量子化回路

- 211 LSPコードブック
- 230 聴感重み付け回路
- 235 減算回路
- 240 応答信号計算回路
- 310 インパルス応答計算回路
- 350、355、356、357 音源量子化回路
- 351、352 音源コードブック
- 370 モード判別回路
- 360 重み付け信号計算回路
- 365、366、367 ゲイン量子化回路
- 380 ゲインコードブック
- 400 マルチプレクサ
- 500 適応コードブック回路
- 510 デマルチプレクサ
- 510 ゲイン復号回路
- 520 適応コードブック回路
- 530 モード判別回路
- 540 音源信号復元回路
- 550 加算回路
- 560 合成フィルタ回路
- 570 スペクトルパラメータ復号回路
- 600 乱数発生回路
- 800 モード判別回路

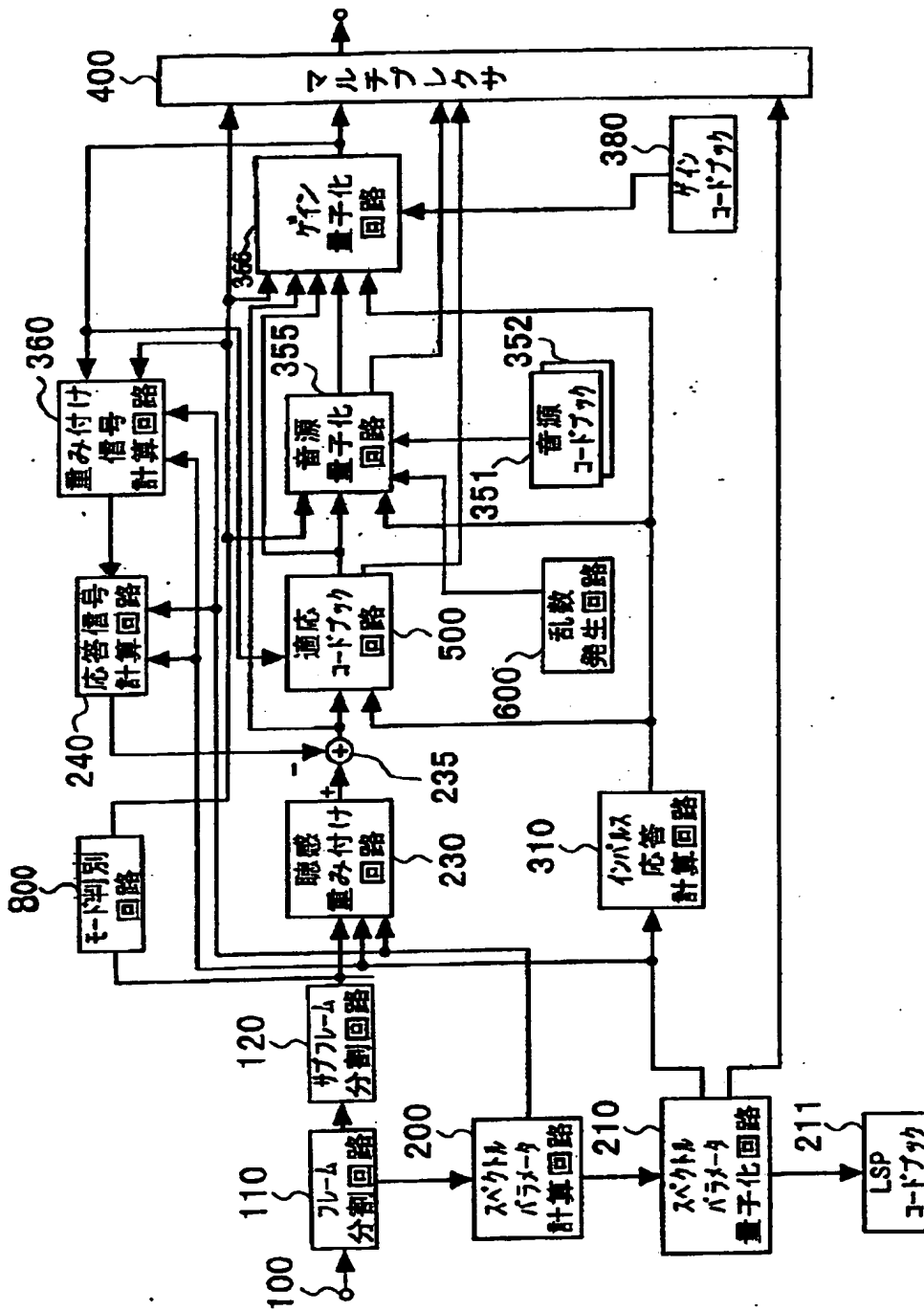
【書類名】

図面

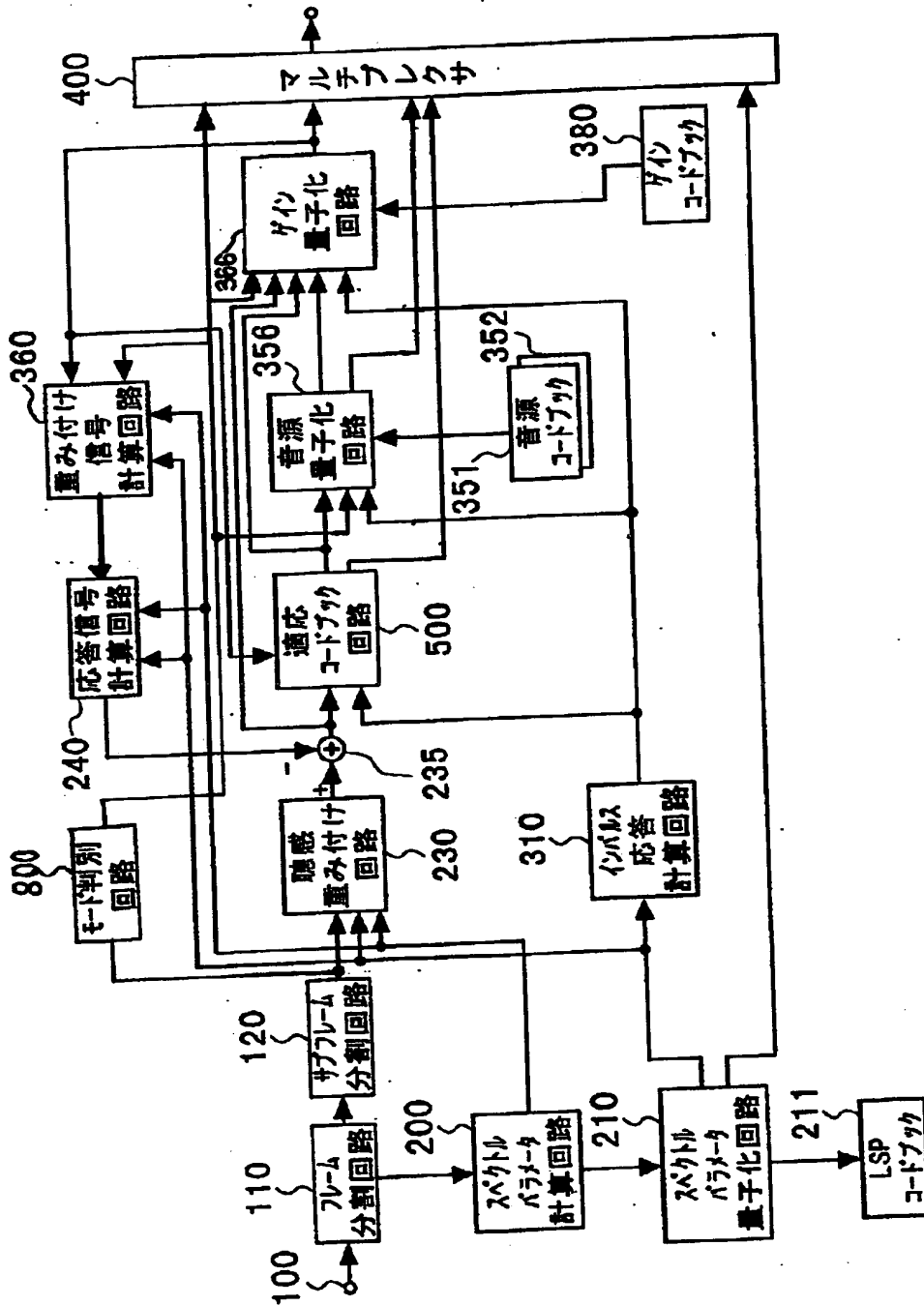
【図1】



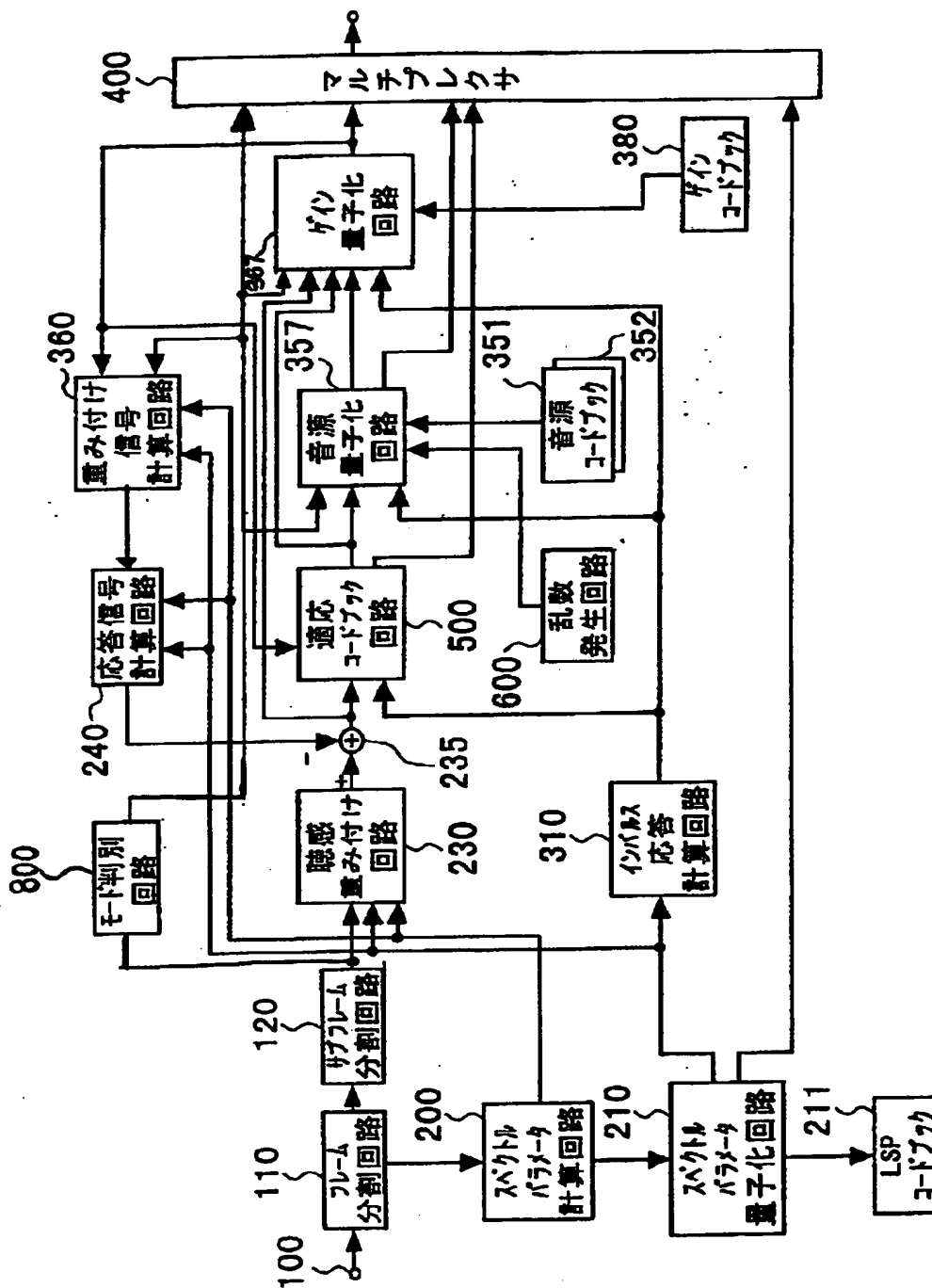
【図 2】



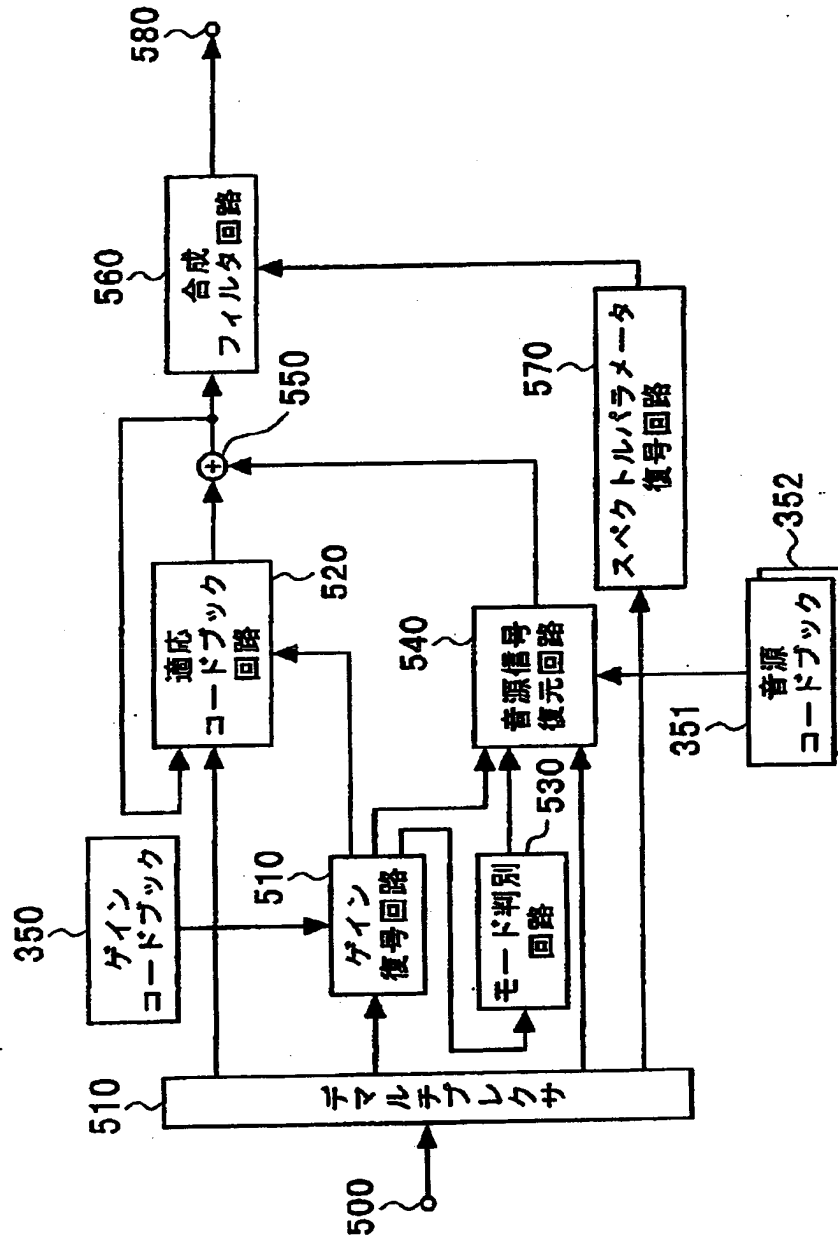
【図 3】



【図 4】



【図5】



【書類名】 要約書

【要約】

【課題】

低ビットレートでも良好な音質の得られる音声符号化装置の提供。

【解決手段】

音声符号化装置のモード判別回路 800 において、サブフレーム毎に入力音声信号から特徴量を用いてモードを判別し、音源量子化回路 350 において予め定められたモードの場合複数個のパルスの振幅もしくは極性をまとめて量子化するコードブック 351、352 に格納されるコードベクトルと、予め定められたパルスの位置を時間的にシフトする複数のシフト量の各々とを組み合わせで探索し、入力音声との歪みを最小にするコードベクトルとシフト量の組合せを選択し、ゲイン量子化回路 365 においてゲインコードブック 380 を用いてゲインを量子化する。

【選択図】

図 1

【書類名】

職権訂正データ

【訂正書類】

特許願

<認定情報・付加情報>

【特許出願人】

【識別番号】

000004237

【住所又は居所】

東京都港区芝五丁目7番1号

【氏名又は名称】

日本電気株式会社

【代理人】

申請人

【識別番号】

100080816

【住所又は居所】

神奈川県横浜市港北区新横浜3丁目20番12号

望星ビル7階加藤内外特許事務所

【氏名又は名称】

加藤 朝道

出 願 人 履 歴 情 報

識別番号 [000004237]

1. 変更年月日	1990年 8月29日
[変更理由]	新規登録
住 所	東京都港区芝五丁目7番1号
氏 名	日本電気株式会社

THIS PAGE BLANK (USPTO)